DISCRETE APPLIED MATHEMATICS

ELSEVIER Discrete Applied Mathematics 57 (1995) 133–144

The complexity of harmonious colouring for trees

Keith Edwards*[,a], Colin McDiarmid[b]

[a] *Department of Mathematics and Computer Science, University of Dundee, Dundee DD1 4HN, UK*
[b] *Department of Statistics, University of Oxford, Oxford OX1 3TG, UK*

Received 4 August 1992; revised 31 August 1993

Abstract

A *harmonious colouring* of a simple graph G is a proper vertex colouring such that each pair of colours appears together on at most one edge. The *harmonious chromatic number* $h(G)$ is the least number of colours in such a colouring. It was shown by Hopcroft and Krishnamoorthy (1983) that the problem of determining the harmonious chromatic number of a graph is NP-hard. We show here that the problem remains hard even when restricted to trees.

1. Introduction

A *harmonious colouring of a simple graph* G is a proper vertex colouring such that each pair of colours appears together on at most one edge. Formally a harmonious colouring is a function c from a colour set C to the set $V(G)$ of vertices of G such that for any edge e of G, with endpoints x, y say, $c(x) \neq c(y)$, and for any pair of distinct edges e, e', with endpoints x, y and x', y', respectively, then $\{c(x), c(y)\} \neq \{c(x'), c(y')\}$. The *harmonious chromatic number* $h(G)$ is the least number of colours in such a colouring. A recent survey on harmonious colourings is by Wilson [5].

It was shown by Hopcroft and Krishnamoorthy [4] that the problem of determining the harmonious chromatic number of a graph is NP-hard, and a short proof of the same result, due to D.S. Johnson, appears in the same paper. In this paper, we show that determining the harmonious chromatic number of a tree is NP-hard. Garey and Johnson [3] list only a few other examples of natural problems which are NP-complete for trees, such as BANDWIDTH, SUBGRAPH ISOMORPHISM and WEIGHTED DIAMETER.

In Section 2 we prove a technical lemma on harmonious colourings of graphs which consist of a disjoint union of stars (a "forest of stars"), and use this lemma to prove that determining the harmonious chromatic number of a tree is NP-hard. In Section 3

* Corresponding author.

SSDI 0166-218X(94)00100-6

we use the technical lemma again to give a formula for the harmonious chromatic number of a forest of stars, and derive some corollaries.

2. Trees

We start with a technical lemma. For any subsets X, Y of the vertices of a graph G, denote by $E(X)$ the set of edges in G which join two vertices in X, and by $E(X, Y)$ the set of edges which join a vertex in X and a vertex in Y.

Lemma 2.1. *Let $G = (V, E)$ be an undirected graph, and for each $v \in V$, let $a(v)$ be a non-negative integer. Then it is possible to orient the edges of G so that for each $v \in V$, the outdegree $d^+(v)$ is at least $a(v)$, if and only if*

$$\sum_{x \in X} a(x) \leqslant |E(X)| + |E(X, V - X)| \quad \textit{for each } X \subseteq V.$$

Remarks. The above result is perhaps "folk-knowledge". We may deduce from it (or from its proof) the complementary result, that we can orient G so that $d^+(v) \leqslant a(v)$ for each vertex $v \in V$ if and only if $|E(X)| \leqslant \sum_{x \in X} a(x)$ for each $X \subseteq V$.

Also, the same proof idea yields the following "interpolation" result. If we have integers $a(v) \leqslant b(v)$ for each vertex $v \in V$ then we can orient G so that $a(v) \leqslant d^+(v) \leqslant b(v)$ for each $v \in V$ if and only if we can orient G to respect the lower bounds (i.e. $a(v) \leqslant d^+(v)$ for each $v \in V$) and we can orient G to respect the upper bounds.

Proof. The condition is clearly necessary, since for any $X \subseteq V$, we have

$$\sum_{x \in X} d^+(x) \leqslant |E(X)| + |E(X, V - X)|.$$

For the converse, suppose that the condition holds, and consider an orientation which minimises the "deficiency" given by

$$\sum_{v \in V} (a(v) - d^+(v))^+.$$

If the deficiency is 0, then the result holds, so suppose that it is positive. Then there must be a vertex v with $a(v) > d^+(v)$, hence the set

$$A = \{v \in V \mid a(v) > d^+(v)\}$$

is non-empty. Then also the set

$$B = \{v \in V \mid a(v) < d^+(v)\}$$

is non-empty, since

$$\sum_{v \in V} a(v) \leqslant |E(V)| = \sum_{v \in V} d^+(v).$$

Now there cannot be a directed path from B to A, for if so we could reverse the orientations on this path to obtain an orientation with smaller deficiency. Let X be the set of vertices from which a vertex of A is reachable along a directed path. Since there are no edges directed from $V - X$ to X, and X and B are disjoint, we have

$$\sum_{x \in X} a(x) > \sum_{x \in X} d^+(x) = |E(X)| + |E(X, V - X)|,$$

a contradiction. The result follows. $\square$

The next lemma gives a condition for a forest of stars to be colourable in a particular way. We will use this lemma in the proof of NP-completeness later in this section.

Lemma 2.2. *Let F be a forest consisting of t stars $F_1, \ldots, F_t$ with sizes (i.e. number of edges) $m_1 \geqslant m_2 \geqslant \cdots \geqslant m_t$, respectively, and let $C \geqslant t$. Then F can be coloured harmoniously with C colours so that the centres of the stars all receive distinct colours if and only if*

$$\sum_{i=1}^{k} m_i \leqslant \sum_{i=1}^{k} (C - i) \quad \textit{for each } k = 1, \ldots, t.$$

Proof. First suppose that F is harmoniously coloured with C colours such that the centres of the stars receive distinct colours. By permuting the colours if necessary, we can assume that the centre of F_i is coloured with colour i. Now consider, for each k, the graph consisting of $F_1, \ldots, F_k$. At most $k(k-1)/2$ edges in this graph have both ends with colours from $1, 2, \ldots, k$, since the colouring is harmonious, and there are at most $k(C - k)$ other edges. Also the total number of edges is $\sum_{i=1}^{k} m_i$. Hence

$$\begin{aligned}\sum_{i=1}^{k} m_i &\leqslant k(k-1)/2 + kC - k^2 \\ &= kC - k(k+1)/2 \\ &= \sum_{i=1}^{k} (C - i).\end{aligned}$$

For the converse, consider the complete graph K_C on vertex set $V = \{v_1, \ldots, v_C\}$. Set $m_i = 0$ for any i with $t < i \leqslant C$. By the previous lemma, K_C can be oriented such that the outdegree of v_i is at least m_i, for $i = 1, \ldots, C$, if for each subset X of V,

$$\sum_{v_i \in X} m_i \leqslant |E(X)| + |E(X, V - X)|.$$

Let X be any subset of V and let $k = |X|$. Then

$$|E(X)| + |E(X, V - X)| = k(k-1)/2 + k(C - k) = \sum_{i=1}^{k} (C - i).$$

Hence since $m_1 \geqslant \cdots \geqslant m_C$, we have

$$\begin{aligned}\sum_{v_i \in X} m_i &\leqslant \sum_{i=1}^{k} m_i \\ &\leqslant \sum_{i=1}^{k} (C-i) \\ &= |E(X)| + |E(X, V-X)|\end{aligned}$$

as required.

Now to colour F, use C colours $1, \ldots, C$. Colour the centre of F_i with colour i and colour the m_i leaves of F_i with m_i colours j such that edge (v_i, v_j) is oriented away from v_i. It is clear that this gives a harmonious colouring of F. $\square$

Definition. The *radius* of a graph G is the minimum over all vertices r of the maximum value of the distance $d(r, v)$ between r and any vertex v.

Corollary 2.3. *Let T be a tree of radius at most 2, so that there is a vertex r such that $d(v,r) \leqslant 2$ for all $v \in V(T)$. Suppose that vertex r has t neighbours, with degrees $m_1 \geqslant \cdots \geqslant m_t$. Then T can be harmoniously coloured with C colours if and only if $C \geqslant t+1$ and for each $k = 1, \ldots, t$*

$$\sum_{i=1}^{k} m_i \leqslant \sum_{i=1}^{k} (C-i).$$

Proof. Firstly, suppose that T can be harmoniously coloured with C colours. Then $C \geqslant t+1$ since vertex v has degree t. Further, by Lemma 2.2 the other conditions must hold, since the t stars of sizes $m_1, \ldots, m_t$ centred on the neighbours of r are harmoniously coloured with C colours with the centres receiving distinct colours.

Conversely, suppose the conditions hold. Then by Lemma 2.2, the stars centred at the neighbours of r can be harmoniously coloured so that the centres receive distinct colours. There is a colour c say not used on any of these centres, so we can assume that this colour is used on some leaf of each of the stars. These leaves coloured c can then be identified to form r. $\square$

We now prove that the problem of determining the harmonious chromatic number of a tree is NP-complete. To make the exposition clearer, it helps to introduce an intermediate problem which we will call problem Π:

Instance: Integer C, and forest F consisting of $3C$ non-trivial stars, partitioned into 3 C-element sets S_W, S_X and S_Y.

Question: Is there a harmonious colouring of F with C colours in which, for each $i = 1, \ldots, C$, exactly 3 of the stars have centre coloured i, one each from S_W, S_X and S_Y?

Lemma 2.4. *The problem* Π *is* NP-*complete.*

Proof. The problem is obviously in NP. To prove completeness, we reduce from the NP-complete problem NUMERICAL 3-DIMENSIONAL MATCHING [3]:

Instance: Disjoint sets W, X and Y, each containing m elements, a size $s(a) \in Z^+$ for each element $a \in W \cup X \cup Y$, and a bound $B \in Z^+$.

Question: Can $W \cup X \cup Y$ be partitioned into m disjoint sets $A_1, \dots, A_m$, such that each A_i contains exactly one element from each of W, X and Y and such that, for $1 \leqslant i \leqslant m$, $\sum_{a \in A_i} s(a) = B$?

Note that NUMERICAL 3-DIMENSIONAL MATCHING is NP-complete in the strong sense [3], thus we can restrict our attention to instances for which B and each $s(a)$ is at most $p(m)$ for some fixed polynomial p. It is helpful to assume that m is odd; this is easy since we can just add new elements with sizes $B - 2, 1, 1$ to W, X and Y if necessary. Also note that we can assume that each $s(a) \leqslant B$ and that each $s(a)$, and B, is a multiple of m, for otherwise we can replace each $s(a)$ by $ms(a)$ and B by mB. Finally we assume that the sum of the sizes of all the elements in $W \cup X \cup Y$ is mB. Let $W = \{w_1, \dots, w_m\}$, and similarly for X and Y.

Now given an instance I of NUMERICAL 3-DIMENSIONAL MATCHING, satisfying these constraints, we construct an instance I' of the problem Π as follows: The constant C is set to $6B + m + 7$. The forest F is formed of the 3 sets S_W, S_X, S_Y, each containing C stars. S_X and S_Y each contain one star S_a of size $s(a)$ for each a in X and Y, respectively, and $C - m$ stars of size $B + 1$. S_W contains a star S_a of size $s(a) + C - \frac{1}{2}(m + 1) - B$ for each $a \in W$, and $C - m$ stars of size $B + 1$. (Recall that m is assumed to be odd.) Note that I' is easily constructed from I in polynomial time.

We now claim that the forest F can be harmoniously coloured with C colours with one star from each of S_W, S_X, S_Y having centre coloured i for each i if and only if I has a solution. So suppose that F has such a harmonious colouring. The centres of the stars in set S_W are coloured $1, \dots, C$ in some order; hence without loss of generality we can assume that the star S_{w_i} is coloured i for $i = 1, \dots, m$. Now the total degree of the centres coloured $1, \dots, m$ is at most

$$\binom{m}{2} + m(C - m) = m(C - \tfrac{1}{2}(m + 1)),$$

while the total degree of the centres of S_{w_i}, $i = 1, \dots, m$, is

$$m(C - \tfrac{1}{2}(m + 1) - B) + \sum_{i=1}^{m} s(w_i).$$

Hence the total size of the stars from S_X and S_Y with centres coloured $1, \dots, m$ is at most

$$mB - \sum_{i=1}^{m} s(w_i).$$

Since the total size of the objects in $X \cup Y$ is exactly this number, and all of the stars not of the form S_a have size $B + 1$ which is greater than any element of $W \cup X \cup Y$, then the only way in which this is possible is if the stars S_{x_i}, $i = 1, \dots, m$, have centres coloured $1, \dots, m$ in some order and similarly for the stars S_{y_i}. But then let

$$A_i = \{a \mid \text{centre of } S_a \text{ has colour } i\}.$$

Then for each $i = 1, \dots, m$, A_i contains exactly one element from each of W, X and Y, and since the neighbours of the centres coloured i all have distinct colours, then

$$C - \tfrac{1}{2}(m + 1) - B + \sum_{a \in A_i} s(a) \leqslant C - 1,$$

hence

$$\sum_{a \in A_i} s(a) \leqslant B - 1 + \tfrac{1}{2}(m + 1).$$

But since each $s(a)$ and B are divisible by m, then the sum on the left must in fact be at most B, and since the sum of all the elements in $W \cup X \cup Y$ is mB, the sums must be equal to B. Hence the instance I of NUMERICAL 3-DIMENSIONAL MATCHING has a solution.

Conversely, suppose that we have a solution $A_1, \dots, A_m$ to the instance I. Then let the forest F_i consist of the stars S_a, $a \in A_i$, for each $i = 1, \dots, m$, and consist of one star of size $B + 1$ from each of S_W, S_X and S_Y for $i = m + 1, \dots, C$. Let the size of F_i be M_i, so that

$$M_i = C - \tfrac{1}{2}(m + 1)$$

for $i = 1, \dots, m$ and $M_i = 3(B + 1) = \tfrac{1}{2}(C - m - 1)$ otherwise. Then by Lemma 2.2, F can be harmoniously coloured with C colours such that all of the centres in F_i have colour i, provided that, for each $k = 1, \dots, C$,

$$\sum_{i=1}^{k} M_i \leqslant \sum_{i=1}^{k} (C - i).$$

Now if $k \leqslant m$, then

$$\begin{aligned}\sum_{i=1}^{k} M_i &= k(C - \tfrac{1}{2}(m + 1)) \\ &= kC - \tfrac{1}{2}k(m + 1) \\ &\leqslant kC - \tfrac{1}{2}k(k + 1) \\ &= \sum_{i=1}^{k} (C - i),\end{aligned}$$

with equality when $k = m$. If $m < k \leqslant C$, then

$$\sum_{i=1}^{k} M_i - \sum_{i=1}^{k} (C - i) = \sum_{i=m+1}^{k} M_i - \sum_{i=m+1}^{k} (C - i)$$
$$= \tfrac{1}{2}(C - m - 1)(k - m) - \tfrac{1}{2}(k - m)(2C - k - m - 1)$$
$$= \tfrac{1}{2}(k - m)(k - C) \leqslant 0.$$

The result follows. □

Proposition 2.5. *The following problem is* NP-*complete.*
Instance: *Forest F consisting of three trees of radius two, integer k.*
Question: *Is $h(F) \leqslant k$?*

Proof. The problem is obviously in NP. To prove completeness we reduce from the problem Π which is proved NP-complete in Lemma 2.4. So let F, consisting of $3C$ stars in 3 C-element sets S_W, S_X and S_Y, and integer C, be an instance of Π. Then construct a forest F' as follows: Take 3 new vertices r_W, r_X and r_Y, and join r_W to the centre of each of the stars in S_W, r_X to the centre of each of the stars in S_X, and r_Y to the centre of each of the stars in S_Y. Take 3 further new vertices and join 2 of them to r_W, and the other to r_X. Hence F consists of 3 trees of depth 2. Take $k = C + 3$.

Now suppose that F' has a harmonious colouring with k colours. Then clearly r_W, r_X and r_Y have distinct colours, and we may assume that these are $C + 3$, $C + 2$ and $C + 1$, respectively. Since these three vertices have degrees $C + 2$, $C + 1$ and C, respectively, then all the colour pairs containing one of these three colours are used on the edges incident with r_W, r_X and r_Y, and no other colour pairs are used. No centre of a star in S_W, S_X or S_Y can therefore have one of the colours $C + 3$, $C + 2$ and $C + 1$, and furthermore no two centres in one of S_W, S_X and S_Y can have the same colour. Hence we obtain a solution to the problem Π.

Conversely, given a colouring of F in which, for each i, one centre from each of S_W, S_X and S_Y is coloured i, then we can colour r_W with colour $C + 3$ and its two extra neighbours $C + 2$ and $C + 1$, colour r_X and its extra neighbour with colours $C + 2$ and $C + 1$, respectively, and colour r_Y with colour $C + 1$. This completes the proof. □

Theorem 2.6. *The following problem is* NP-*complete.*
Instance: *Tree T of radius* 3, *integer k.*
Question: *Is $h(T) \leqslant k$?*

Proof. The problem is obviously in NP. The proof of completeness is identical to that of Proposition 2.5 except that we identify one of the extra neighbours of r_W with r_X and the other with r_Y. □

Remark. From Corollary 2.3 we see that if T is a tree of radius at most 2 then we can determine $h(T)$ in polynomial time. Proposition 2.5 and Theorem 2.6 above show that it is NP-hard to determine $h(T)$ if T is a forest consisting of 3 trees of radius 2, or if T is a tree of radius 3.

3. Forests of stars

In this section we consider further the problem of finding harmonious colourings of forests of stars, using Lemma 2.2 from the previous section. We first need another technical lemma.

Lemma 3.1. *Let $m_1 \geqslant m_2 \geqslant \cdots \geqslant m_t$ be positive integers and let C be a positive integer. Then we can partition the set of integers $m_1, \ldots, m_t$ into C parts with sums $M_1 \geqslant \cdots \geqslant M_C$, such that $\sum_{i=1}^{k} M_i \leqslant \sum_{i=1}^{k} (C-i)$ for each $k = 1, \ldots, C$, if and only if*

$$\sum_{i=1}^{k} m_i \leqslant \sum_{i=1}^{k} (C-i) \quad \text{for each } k = 1, \ldots, C$$

and

$$\sum_{i=1}^{t} m_i \leqslant \binom{C}{2}.$$

Proof. It is easy to see that the condition is necessary. To prove that it is sufficient, we apply the natural heuristic for scheduling t jobs with processing times m_i on C parallel machines, namely allocating the jobs sequentially to the first available machine.

Let $x_1, \ldots, x_C$ be integers initially all 0. We work through the m_i's in order, adding each in turn to x_C, then reordering the x's so that $x_1 \geqslant \cdots \geqslant x_C$. We claim that after each stage, for each k,

$$\sum_{i=1}^{k} x_i \leqslant \sum_{i=1}^{k} (C-i).$$

Thus when we have used all of the m's we take $M_i = x_i$.

To establish the claim, for the first C stages we get, after reordering, $x_1 = m_1$, $x_2 = m_2$, and so on, so the claim for these stages follows immediately from the hypotheses. So suppose that $i > C$, and we wish to add m_i. Suppose the current value of x_C is a. Then $m_i \leqslant a$, since a consists of at least one m_j with $j < i$. Adding m_i forms $m_i + a = b$. Reordering gives a new ordering $x'_1 \geqslant \cdots \geqslant x'_C$, with $b = x'_j$ say, where $1 \leqslant j \leqslant C$. Now suppose that, contrary to the claim, there is a k such that

$$\sum_{i=1}^{k} x'_i > \sum_{i=1}^{k} (C-i),$$

and that this is the first time that the claim fails. Then clearly we must have $k \geqslant j$. So

$$\begin{aligned} \sum_{i=1}^{k} (C-i) < \sum_{i=1}^{k} x'_i &= \sum_{i=1}^{k-1} x_i + b \\ &\leqslant \sum_{i=1}^{k-1} (C-i) + b. \end{aligned}$$

Hence $b > C - k$. But also $a \geqslant m_i$, so $2a \geqslant m_i + a = b$, or $a > \frac{1}{2}(C - k)$. Now

$$\begin{aligned}\sum_{i=1}^{C} x_i' &\geqslant \sum_{i=1}^{k} x_i' + (C-k)x_C' \\ &\geqslant \sum_{i=1}^{k} x_i' + (C-k)a \quad (\text{since } x_C' \geqslant x_C = a) \\ &> \sum_{i=1}^{k} (C-i) + \tfrac{1}{2}(C-k)^2 \\ &\geqslant \binom{C}{2},\end{aligned}$$

a contradiction since clearly $\sum_{i=1}^{t} m_i \geqslant \sum_{i=1}^{C} x_i'$. $\square$

We now prove the main theorem of Section 3 which gives easily checked conditions for a forest of stars to be harmoniously colourable with C colours.

Theorem 3.2. *Let F be a forest consisting of t stars of sizes $m_1 \geqslant \cdots \geqslant m_t$. Then F can be coloured harmoniously with C colours if and only if*

$$\sum_{i=1}^{k} m_i \leqslant \sum_{i=1}^{k} (C-i) \quad \textit{for each } k = 1, \ldots, C,$$

and

$$\sum_{i=1}^{t} m_i \leqslant \binom{C}{2}.$$

Proof. First suppose that F is harmoniously coloured with C colours. For each $k = 1, \ldots, C$, let F_k be the subforest consisting of the k largest stars, so that F_k has $\sum_{i=1}^{k} m_i$ edges. Let A be the set of colours used on the centres of the stars and let $k' = |A|$. Then every edge of F_k has either (i) both ends coloured with a colour from A; there are at most $k'(k'-1)/2$ such edges, or (ii) exactly one end coloured with a colour from A; there are at most $k'(C-k')$ such edges. Hence

$$\begin{aligned}\sum_{i=1}^{k} m_i &\leqslant k'(k'-1)/2 + k'C - k'^2 \\ &= k'C - k'(k'+1)/2 \\ &= \sum_{i=1}^{k'} (C-i) \\ &\leqslant \sum_{i=1}^{k} (C-i).\end{aligned}$$

For the converse result, suppose now that the conditions hold. Then, by Lemma 3.1, we can partition F into C subforests $F_1, \ldots, F_C$, with $M_1, \ldots, M_k$ edges, respectively, such that $\sum_{i=1}^{k} M_i \leqslant \sum_{i=1}^{k} (C-i)$ for each $k = 1, \ldots, C$. If, for each $i = 1, \ldots, C$, we identify the centres of all the stars in F_i, then we obtain a forest of C stars, and, by Lemma 2.2, this forest can be coloured with C colours. Splitting each star into the corresponding forest F_i immediately gives a colouring of F. □

Definition. For any positive integer m, let $Q = Q(m)$ be the least integer k such that $\binom{k}{2} \geqslant m$. It is easily checked that

$$Q(m) = \left\lceil \tfrac{1}{2}(\sqrt{1+8m}+1) \right\rceil.$$

Corollary 3.3. *Let F be a forest consisting of t stars of sizes $m_1 \geqslant \cdots \geqslant m_t$. Let $m = \sum_{i=1}^{t} m_i$, the number of edges of F. Then*

$$h(F) = \max\left(\max\left\{\left\lceil \frac{1}{k}\sum_{i=1}^{k}(m_i+i)\right\rceil \,\middle|\, 1 \leqslant k < Q(m)\right\}, Q(m)\right).$$

Proof. From Theorem 3.2, $h(F)$ is the least integer $C \geqslant Q(m)$ such that

$$\sum_{i=1}^{k} m_i \leqslant \sum_{i=1}^{k} (C-i) \quad \text{for each } k = 1, \ldots, C.$$

If C and k are at least Q, then

$$\sum_{i=1}^{k} (C-i) \geqslant \sum_{i=1}^{Q} (Q-i) = \binom{Q}{2} \geqslant m \geqslant \sum_{i=1}^{k} m_i,$$

hence the conditions for $k \geqslant Q$ are redundant. Rearranging the remaining conditions gives the formula. □

If a graph G has m edges then $h(G) \geqslant Q(m)$. It is of particular interest to investigate when $h(G) = Q$ or $h(G)$ is very near to Q (see for example [1]). The following corollary gives an example of this.

Corollary 3.4. *Let F be a forest of stars, let m be the number of edges of F, and let Δ be the maximum degree of F. If $\Delta \leqslant (\sqrt{1+8m}-1)/4$, then $h(F) = Q(m)$.*

Proof. From the condition, it is easily calculated that $\Delta \leqslant (Q-1)/2$. Then if $1 \leqslant k < Q(m)$, we have

$$\frac{1}{k}\sum_{i=1}^{k}(m_i+i) \leqslant \frac{Q-1}{2} + \frac{k+1}{2} < Q.$$

The result follows from Corollary 3.3. □

Corollary 3.5. *Let F be a forest consisting of a_i copies of a b_i-star, for $i = 1, \ldots, t$, where $b_1 > \cdots > b_t$. Then F can be coloured with C colours harmoniously if and only if, for each j such that $a_1 + \cdots + a_j \leqslant C$,*

$$\sum_{i=1}^{j} a_i b_i \leqslant \sum_{i=1}^{a_1+\cdots+a_j} (C-i)$$

and

$$\sum_{i=1}^{t} a_i b_i \leqslant \binom{C}{2}.$$

Proof. We have to ensure, from Theorem 3.2, that

$$\sum_{i=1}^{k} m_i \leqslant \sum_{i=1}^{k} (C-i)$$

for each $k = 1, \ldots, C$, where the m_i are the sizes of all the component stars in decreasing order. Since $\sum_{i=1}^{k} (C-i)$ is convex as a function of k, while $\sum_{i=1}^{k} m_i$ is piecewise linear, we only need check the inequality when m_i changes value from one b_j to the next. The result follows. □

Corollary 3.6. *The following problem can be solved in polynomial time:*

Instance: Positive integers $a_1, \ldots, a_t, b_1, \ldots, b_t, C$, given in binary.

Question: Can the forest consisting of a_i copies of a b_i-star, for $i = 1, \ldots, t$, be harmoniously coloured with C colours?

Proof. This is clear from Corollary 3.5. □

4. Line-distinguishing colourings

A *line-distinguishing colouring* of a graph is like a harmonious colouring except that it need not be proper. Thus for each colour there is at most one edge with both ends of that colour. The *line-distinguishing number* $ld(G)$ is the least number of colours in such a colouring. For all of our results on harmonious colouring there are analogous results on line-distinguishing colourings. To see this, note first the following observation which allows us to use Lemma 2.2.

Lemma 4.1. *Let F be a forest consisting of t stars of sizes $m_1, \ldots, m_t \geqslant 1$ and let $C \geqslant t$. Then F has a line-distinguishing colouring with C colours so that the centres of the stars all receive distinct colours if and only if the forest F' has a similar harmonious colouring, where F' consists of stars of sizes $m_1 - 1, \ldots, m_t - 1$.*

Corresponding to Corollary 2.3 we have the following proposition.

Proposition 4.2. *Let T be a tree of radius at most 2, so that there is a vertex r such that $d(v,r) \leqslant 2$ for all $v \in V(T)$. Suppose that vertex r has t neighbours with degrees $m_1 \geqslant \cdots \geqslant m_t \geqslant 2$. Then T has a line-distinguishing colouring with C colours if and only if $C \geqslant t+1$ and, for each $k = 1, \ldots, t$,*

$$\sum_{i=1}^{k} m_i \leqslant kC - \binom{k}{2}.$$

Hence for trees T of radius at most 2 we can determine $ld(T)$ in polynomial time. However, we need only minor changes to the proof of Theorem 2.6 to prove the following result, which was conjectured by Frank et al. [2].

Theorem 4.3. *The following problem is* NP-*complete.*
Instance: *Tree T of radius* 3, *integer k.*
Question: *Is $ld(T) \leqslant k$?*

5. Concluding remarks

We have seen that it is hard to determine $h(T)$ for trees T. From this we might well suppose that determining the harmonious chromatic number will be hard for any reasonably non-trivial class of graphs. However, the proof of NP-completeness above relies on trees which contain vertices of high degree (around $\sqrt{n}$). It seems likely that if we fix an integer d and restrict our attention to graphs with maximum degree at most d, then the problem may be more tractable at least for trees, anffd perhaps for general graphs. Thus we propose the following conjecture.

Conjecture 5.1. Let d be a fixed positive integer. Then the following problem is solvable in polynomial time:
Instance: Tree T with maximum degree at most d, integer k.
Question: Is $h(T) \leqslant k$?

References

[1] K.J. Edwards and C.J.H. McDiarmid, New upper bounds on harmonious colourings, J. Graph Theory 18 (1994) 257–267.
[2] O. Frank, F. Harary and M. Plantholt, The line-distinguishing chromatic number of a graph, Ars Combin. 14 (1982) 241–252.
[3] M.R. Garey and D.S. Johnson, Computers and Intractability (Freeman, New York, 1980).
[4] J. Hopcroft and M.S. Krishnamoorthy, On the harmonious coloring of graphs, SIAM J. Algebraic Discrete Methods 4 (1983) 306–311.
[5] B. Wilson, Line distinguishing and harmonious colourings, in: R. Nelson and R.J. Wilson, eds., Graph Colourings, Pitman Research Notes in Mathematics 218 (Longman Scientific & Technical, Essex, 1990).